心理科学进展　2018, Vol. 26, No. 6, 1131–1140
Advances in Psychological Science　DOI: 10.3724/SP.J.1042.2018.01131

多源评价的特点与内在机制

张　赟[1]　翁清雄[2]
([1]西南财经大学工商管理学院, 成都 611130) ([2]中国科学技术大学管理学院, 合肥 230026)

摘　要　多源评价在国外企业中的运用已日益成熟，但在我国还停留在探索与发展阶段。基于已有的研究发现，围绕评价过程、评价源及被评价者三方面对多源评价的特点及内在机制进行了探讨与分析。从评价过程看，其评价目的具有多重性，评价形式注重匿名性，且评价结果的合理应用非常重要；从评价源看，不同评价源间的评价一致性较低，且易造成晕轮效应和宽大效应；从被评价者来看，个体对多源评价结果的反应，受到个性特征、反馈信号及自我–他人评价间差距等因素影响。研究也发现，多源评价所带来的绩效改进结果具有不稳定性。基于此，如何提高多源评价过程的有效性与准确性，改善评价者对评价结果的反应，以及如何对多源评价结果进行有效汇总等是未来值得研究的重要内容。

关键词　多源评价；评价一致性；评价偏误
分类号　B849:C93

1　引言

在工业经济时期，在大规模的重复生产方式条件下，对员工的绩效评价主要为上级评价模式(top-down appraisal)。但随着经济的发展，定制生产方式的出现，上级无法拥有员工的完整信息，不能对员工作出客观公正的评价(许庆瑞，王勇，陈劲，2002)，因此拥有多个评价源的评价模式应运而生。多源评价/反馈(multisource feedback)，包括 360 度、270 度或 180 度评价(Markham, Smith, Markham, & Braekkan, 2014)，指被评价者受到来自上级、同事、自我、下级及顾客的多角度评价(London & Smither, 1995)。据统计，全球 500 强企业中，有 90%的企业都在使用多源评价(Maylett, 2009)。相对于传统的上级评价来说，多源评价是多个评价主体同时对被评价者的绩效进行评估，因此能够提供更准确、多角度的描述(Hoffman, Lance, Bynum, & Gentry, 2010)。

目前，我国对多源评价的研究还较少，且大多属于定性研究。如刘凤瑜(2000)对 360 度反馈的基本原理与过程进行了阐述；陆昌勤、方俐洛和凌文辁(2001)探讨了 360 度反馈在人力资源管理中的效用，以及在中国文化背景下的实施策略；许庆瑞等(2002)对多源评价中不同的评价源特征及缺点进行了分析，并针对不同的组织环境对多源评价方式进行了匹配。然而，目前国内关于多源评价的定量研究，却几乎没有；这与国外的研究文献产生了很大的反差。本文基于国外关于多源评价的实证研究，从评价过程、评价源[1]及被评价者三方面，对国外多源评价的特点及内在机制进行探讨，旨在总结现有研究的不足，并提出未来的研究方向。针对内在机制，本文主要围绕评价结果，分析评价过程、评价源对于评价结果的影响，评价者对评价结果反应的影响因素，以及各因素间的相互作用，如图 1 所示。

2　多源评价过程

2.1　评价目的

多源评价最初主要以帮助员工自身发展为目的(London & Smither, 1995)，通过识别员工的优缺点，制定相应的绩效改进计划，以促进员工发展。如 Wong 和 Kwong (2007)提出，多源评价应

收稿日期：2016-11-16
通信作者：翁清雄, E-mail: wqx886@ustc.edu.cn

[1] 本文中的评价源指代同一类别评价者的集合，而评价者则指代单个的评价人。

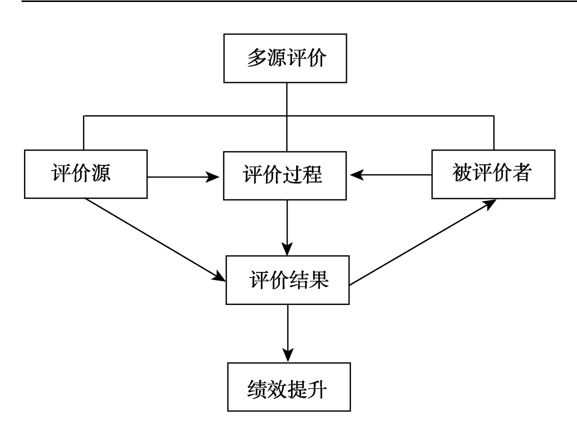


图 1 多源评价的结构图

以追求和谐、追求公平、识别员工优缺点及激励员工为目的。随着多源评价在实际运用中的不断发展，其评价目的也变得更为丰富。比如，不少企业开始将多源评价运用于管理目的。London (2001)认为多源评价既适用于发展目的，也可用于管理目的，但若运用到管理活动中则需花费更多时间。当多源评价以员工个人发展为目的时，会减少评价者和被评价者的顾虑，评价者会作出更诚实的评价，评价结果也更容易被接受(Rogers, Rogers, & Metlay, 2002)。而当多源评价运用于管理目的时，由于评价结果与员工的奖金、晋升等相关，为了追求和谐的组织氛围，评价者往往会提高对员工的评价，从而更易导致宽大效应的出现(Chappelow, 2004)。有趣的是，最近研究发现当多源评价同时用于管理与发展目的时，被评价者由于受到两大目的的同时激励，相对于单一评价目的来说，评价结果对员工知识共享能力的影响将增强(Kim, Atwater, Patel, & Smither, 2016)。可见，多源评价的评价目的由单一性向多重性发展，企业在运用多源评价时，应根据自身特点及发展需求，设立多重而清晰的多源评价目的。

2.2 评价形式

在多源评价过程中一般采用匿名的形式进行评价，因为当评价者得知个人信息不会被泄露时，他们会作出更真实而准确的评价(Brutus & Derayeh, 2002)。特别是下级评价中，当下级员工给与管理者较低评分时，常担心上级领导的报复行为，从而导致他们提高对管理者的评分(Atwater, Brett, & Charles, 2007)。因此，多源评价过程应坚持匿名性原则，以保证评价结果的准确性。同时，评价量表的制定也对评价结果有一定的影响。多源评价过程中，一般以李克特量表的形式，让评价者进行打分。但 Brown, Inceoglu 和 Lin (2017)提出，李克特量表会带来较强的反应偏差(response biases)，而采用强迫排序(forced-choice ranking)来替代李克特量表，能降低评价误差，从而提高评价者间的一致性。可见，从改进评价量表形式的角度入手，来提高评价者间的一致性，是未来研究值得探讨的问题。另外，在反馈结果的呈现上，Atwater 等(2007)的研究结果显示，相对于文字型的反馈结果来说，被评价者更喜欢数字型的反馈，因为文字型的反馈常表达得含糊不清，而数字型的反馈则更为具体而易懂。

2.3 评价后的引导机制

有研究发现，在多源评价结束几个月之后，大多员工将记不起绩效评价的结果，特别是关于其弱点方面的内容(Smither, Brett, & Atwater, 2008)。而此时若配一名优秀的引导者(如有经验的老员工)，将会提高员工对反馈结果的感知有用性及自我意识，降低其对负面反馈的消极情绪，增加其改进动力，从而最终获得绩效提升(Luthans & Peterson, 2003)。另外，相关学者还发现上级的领导风格或行为，会影响员工的后续绩效改进行为。如教练式领导风格能负向显著地影响，员工对绩效评价中组织政治(organizational politcis)的感知，即绩效评价过程是否公正严明或受到其他政治因素的影响(Russo, Miraglia, & Borgogni, 2017)；而领导的自主支持行为，会促进员工的反馈寻求行为，让员工更主动积极的，根据反馈结果来提高个人绩效(Beenen, Pichler, & Levy, 2017)。因而，在员工收到多源评价结果后，应尽量为其分配一名优秀的引导者，帮助其更好地利用评价结果，来获得自身能力的提高。同时，领导者也应以身作则，通过树立良好的领导风格与行为，潜移默化地影响员工的后续绩效改进活动。

3 多源评价中的评价源

相对于传统的绩效评价，多源评价的最大特点就是拥有不同的评价源，关于顾客这一评价源的研究较少，因此本文涉及的评价源主要包括上级、同级、下级及被评价者本身。

3.1 评价源间的评价一致性

虽然已有研究证明了，多源评价中自我评价、同级评价、上级评价及下级评价间的测量等

价性, 即表明不同评价源间的评价结果, 可直接进行比较(Facteau & Craig, 2001; Woehr, Sheehan, & Bennett, 2005)。然而, 由于员工在各类评价源面前, 所表现的行为不同(Braddy, Gooty, Fleenor, & Yammarino, 2014), 各类评价源所关注的重点也有差异(Semeijn, Van Der Heijden, & Van Der Lee, 2014), 因而带来了不同评价源间一致性的问题。以往研究证实了, 多源评价中不同评价源间的评价一致性较低(Hassan & Rohrbaugh, 2009; Beus, Jarrett, Bergman, & Payne, 2012; Braddy et al., 2014)。如 Semeijn 等(2014)认为, 同级评价与下级评价间的分歧最大, 且同类评价源的内部一致性, 要高于评价源间的一致性。同时, 相关学者还从各个方面探讨了其原因, 包括评价者个性特征(Randall & Sharples, 2012), 工作环境氛围(Rosen, Kacmar, Harris, Gavin, & Hochwarter, 2017), 及文化特征(Atwater, Wang, Smither, & Fleenor, 2009; Eckert, Ekelund, Gentry, & Dawson, 2010)等。其中, 文化的影响作用具体表现为, Atwater 等(2009)探讨了国家文化对于自我−他人评价间一致性的调节作用, 并得出在高独断性(assertiveness)及高权力距离(power distance)的文化背景下, 管理者的自我评价与同级评价、下级评价间的一致性相对更高。另外, 评价源间的不一致问题, 会同时影响评价者与被评价者的行为。Amundsen 和 Martinsen (2014)的研究证明, 在对中层管理者进行授权领导行为评价时, 自我评价低于上级评价时, 该管理者的领导效能更高; 当自我评价高于下级评价时, 下级的工作满意度则更低, 离职倾向相对更高。Kopperud, Martinsen 和 Humborstad (2014)证明了, 中层管理者的自我评价低于下级评价时, 组织的服务氛围(service climate)会更高。

在多源评价一致性问题研究中, 自我评价与他人评价间的一致性(self-others rating agreement)问题受到的关注最多(Amundsen & Martinsen, 2014; Kopperud et al., 2014; Markham, Markham, & Smith, 2015)。自我−他人评价一致性, 通常受到性别、年龄及种族等人口统计特征, 个性品质及相关外在环境的影响(Fleenor, Smither, Atwater, Braddy, & Sturm, 2010)。如男性和年老的员工容易高估自己的绩效(Vecchio & Anderson, 2009), 黑人相对于白人来说常为高估者(Bernardin, Konopaske, & Hagan, 2012)。另外, 由于不同评价源间的一致性较低, 以往对评价结果直接平均的做法, 会造成部分重要信息的丢失(Markham et al., 2014)。因此, Markham 等学者基于组内和组间分析(within and between analysis), 利用相关指标, 探讨了多源评价中同一评价源组内平均(Markham et al., 2014)以及不同评价源组间平均, 包括自我−下级评价间(Markham et al., 2015)和自我−同级评价间(Markham, Markham, & Smith, 2017)的平均问题, 并得出不是任何时候都能对多源评价结果, 采用直接平均的方式, 需满足一定的标准条件。可见, 如何提高多源评价中不同评价源间的一致性问题, 是今后多源评价研究的重要方向。

3.2　评价源与评价偏误

晕轮效应(halo effect)和宽大效应(leniency effect), 是多源评价过程中常见的两大评价偏误(rating biases) (Pulakos & O'Leary, 2011), 影响着绩效评价的有效性(Hoffman et al., 2010; Ng, Koh, Ang, Kennedy, & Chan, 2011; O'Neill, Goffin, & Gellatly, 2012) 。其中, 晕轮效应指在被评价者某种特征的影响下, 形成固有的印象, 从而产生以偏概全的认知偏误; 而宽大效应则指评价者在对他人的评价的过程中, 相对于负面评价, 更倾向于正面的肯定与夸赞。

基于对相关文献(Antonioni & Park, 2001; Ng et al., 2011 等)的总结分析, 不同评价源产生评价偏误的概率存在差异。其中, 自我评价过程中出现宽大效应的概率相对最大, 却较少出现晕轮效应。基于归因理论和自我提升理论, 这可能是因为个体为了维持较好的自我感觉, 会扭曲部分信息, 把成功的原因归于自己, 而将失败归结于环境或他人的问题。在下级评价中, 两类评价偏误出现的概率都很大, 且会随着对被评价者观察时间的增加而增强(Antonioni & Park, 2001)。这与下级对上级评价时, 担心其会受到报复行为的心理有关。同级评价出现两类评价偏误的可能性差不多, 常处于选择正面评价或负面评价之间徘徊。因而, 相对于其他评价源来说, 上级评价的结果相对更准确。

另外, 评价偏误受评价者与被评价者间的工作关联性影响。如 Spence 和 Keeping (2010)提出, 评价者的动机会影响其对被评价者的评分, 如上级为了培养下属忠诚度, 会夸大其业绩; 下级为了避免上级的报复, 会故意抬高对其评分; 而同

级之间，由于存在竞争关系，可能会故意压低评分。评价者的自身特点，也会影响评价偏误及评价准确性。如评价者的自我建构(self-construal) (Mishra & Roch, 2013)，人格特征如责任感(Mero, Guidice, & Brownlee, 2007; Bernardin, Tyler, & Villanova, 2009)及随和性(Wood, Harms, & Vazire, 2010; Bono, Hooper, & Yoon, 2012)等因素。Bernardin, Thomason, Buckley 和 Kane (2016)的最新研究发现，随和性较高，自信心较低及绩效管理能力较差的评价者，更易带来评价偏误，降低评价的准确性。同时，文化也会对不同评价源的评价造成影响。比如，个人主义倾向高的管理者，会更容易提高对自我的评价(Xie, Chen, & Roy, 2006)；若组织文化宣扬“合作”而不是“竞争”，则同级评价中故意压低评分的现象将减少(Conway, Lombardo, & Sanders, 2001)。Kossek, Huang, Piszczek, Fleenor 和 Ruderman (2017)证明，评价者与被评价者间的文化距离(culture distance)，即两者的文化差异程度，也会影响评价结果，并受到文化特征的调节作用，具体表现为，在低权力距离和高人性取向的环境下，文化距离对同级评价的负面影响作用，会相对减弱。

4 多源评价中的被评价者

4.1 对评价过程的接受性

个体对多源评价过程的接受意愿，受到自身个性特征的影响。如具有开放精神的员工，会更愿意接受多源评价，并认为其是有价值的(Smither, London, & Reilly, 2005)；自我效能感(self-efficacy)较高的个体，对多源评价持有更积极的态度，并在绩效评价之后获得更多的提升(Heslin & Latham, 2004)。因为，自我效能感较高的个体，会关注如何分析和解决问题，而不是一味地担心失败。Atwater 等(2007)的研究也发现，具有较高自我效能感的个体，在绩效评价之后更愿意与评价者见面，与上级领导讨论，并制定相关的改进目标，同时也会更有动力使用绩效评价结果。而具有犬儒思想的个体，对组织持嘲讽态度时，即不相信组织中存在积极的改变，或认为个人努力在组织中无意义，会更容易拒绝多源评价(McCarthy & Garavan, 2007)。另外，组织文化也会影响个体对多源评价的接受度。London 和 Smither (2002)提出，如果组织内建立了“反馈文化(feedback culture)”，即个体经常接受或请求，获得正式或非正式的反馈来提高自己的工作绩效，个体会更愿意接受多源评价。

4.2 对评价结果的反应

当个体收到多源评价结果后，受各种因素的影响，会表现出不同的反应(reaction)，但学者们对“反应”的相关概念并没有统一的标准(Pichler, 2012; Brown & Benson, 2003)。基于对现有研究的总结分析，本文将个体对评价结果的反应分为三类：感知反应，包括准确性、公平性及有用性；情感反应，包括满意度，积极情绪与消极情绪；行为反应，包括绩效改进动力，设立更具体的目标，与他人分享评价结果，向上级领导求助等。同时，还分析了三大个体反应间的相互关系，如图 2 所示。当被评价者感知评价结果是准确、公平且有效的，会提高其对评价结果的情感反应，表现出较高的满意度及积极的情绪，从而增强其行为反应，具体表现为充满改进动力，设置具体的目标计划，向他人分享结果，并寻求帮助等。

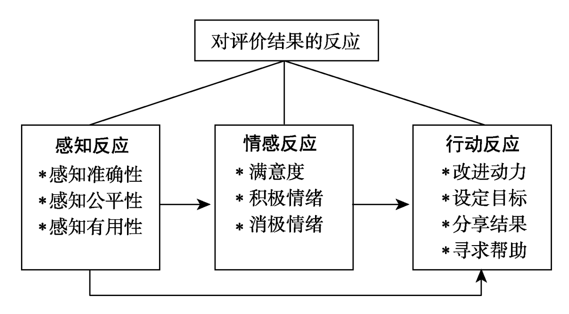


图 2 个体对多源评价评价结果的反应

员工对多源评价结果的反应，受到很多因素的影响，包括评分值(高或低)、个性特征、反馈信号(正面或负面)、自我评价与他人评价间的差距等，并受到个体目标定向、个性特征等的调节作用。其中，绩效评价结果的分值越高，往往被评价者的满意度也越高(Judge & Bono, 2001; Kampkötter, 2017)。然而，Luffarelli, Goncalves 和 Stamatogiannakis (2016)的研究表明，该结果仅针对被评价者的社会比较倾向(social comparison orientation)较低时而成立。被评价者的个性特征的影响作用可表现为，经验开放性较高的员工，对评价结果更易产生较高的满意度(Krasman, 2010)；当评价结果不与奖金相关时，评价结果对员工工作满意度的影响，受到控制点(locus of control)及经验开放性的

负向调节作用(Kampkötter, 2017)。

反馈信号(feedback sign)是指根据反馈内容的特点, 将反馈分为正面反馈和负面反馈两类。被评价者常认为正面反馈比负面反馈, 更加正确和有用(Brett & Atwater, 2001), 从而更乐意接受反馈结果(McCarthy & Garavan, 2007), 表现出较高的满意度及积极的行为反应(Culbertson, Henning, & Payne, 2013), 绩效也更易获得提高(Smither & Walker, 2004)。而负面反馈常被员工认为是不公平, 不准确, 缺乏有用性的(Brett & Atwater, 2001), 他们会选择忽视或者不接受反馈结果, 表现出较低的满意度及消极的情感反应(Culbertson et al., 2013)。但 Brett 和 Atwater (2001)认为, 这只是短暂的情绪反应, 当几周过后, 反馈信号与感知反应间并不存在相关性。同时, Smither 等(2005)发现当被评价者收到差评时, 刚开始会出现消极的反应, 但当 6 个月之后, 他们会设定更多改进目标。另外, 被评价者对正面或负面反馈结果的反应, 还受到个体目标定向的调节作用(Heslin & Latham, 2004; Culbertson et al., 2013)。

自我−他人评价间的差异, 也影响着被评价者对评价结果的反应。当自我评价高于他人评价时, 被评价者常会对反馈结果产生消极的情感反应, 而不同的评价源下的情况不同。当被评价者高估自己时, 会表现出生气、抱怨等负面情绪, 特别是收到较低的同级与上级评价时; 而对下级的负面评价则不易出现消极情绪(Brett & Atwater, 2001), 反而会负向影响下级的行为, 如增加下级的离职倾向(Amundsen & Martinsen, 2014)。当自我评价低于他人评价时, 虽然不一定会带来积极的行为反应(Brett & Atwater, 2001), 但被评价者的满意度会变得更高(Bono & Colbert, 2005), 且受到相关调节变量如员工的核心自我评价(Bono & Colbert, 2005), 评价分值高低(Atwater & Brett, 2005)等的影响。另外, 组织环境也会影响员工对评价结果的反应。如 Miller, Verbos, & Goswami (2014)证明了工作的环境与相关资源、公司沟通交流的开放程度、与同事间的关系及为准备绩效评价所投入的时间, 都会正向影响员工对绩效评价结果的反应。

5　多源评价结果

多源评价的核心目的是通过对员工多方面的评价, 最终提升其绩效。以往研究证明了多源评价后, 员工行为发生了改变, 绩效获得了提升(London & Smither, 1995), 且工作满意度也被提高, 离职率降低, 工作投入感更强(Atwater & Brett, 2006)。Shipper, Hoffman 和 Rotondo (2007)证实了 360 度反馈能够带来认知、情感、行为及产出成果的提升, 但受到国家文化的影响作用, 具体表现为, 在低权力距离及高个人主义的文化背景下, 360 度反馈会相对更有效。然而, 也有学者得出的评价结果不确定(Carson, 2006)。Kim 等(2016)认为, 多源评价对于组织绩效的影响是间接的, 主要通过提高员工的能力与知识共享技能, 从而提高财务业绩。

综上所述, 虽然多源评价在国外企业中的发展已较为成熟, 但其所带来的最终绩效改进效果却并不稳定。基于对现有文献的梳理与总结, 本文以被评价者的“感知反应−情感反应−行为反应−绩效改进结果”为主线, 将多源评价中所涉及的主要因素间的相互关系, 作出了如下图 3 所示的结构图。从图 3 中可看出, 影响员工最终绩效是否能够提升的因素很多, 包括直接效应与间接效应, 如员工对反馈结果的感知反应、情感反应及行为反应的不同, 都会导致不同的绩效改进结果。因此, 绩效改进结果的研究结论出现差异, 也是符合情理的。

6　结论与研究展望

6.1　结论

多源评价已成功的在许多企业中运行, 不仅带来了个人绩效提升, 还能推动组织的进一步发展(Karkoulian, Assaker, & Hallak, 2016)。本文从多源评价过程、评价源及被评价者三个方面, 对国外多源评价的特点及内在机制进行了总结与分析。发现多源评价所带来的绩效改进结果并不稳定, 并受到各种影响因素的直接或间接的影响。从上图 3 中可得, 各变量间的关系错综复杂, 其中个体对反馈结果的情感反应及行为反应, 受到的影响因素最多。而文化特征、个人性格特点及个体目标定向, 是多源评价研究中常用的调节变量。因此, 企业若是在运用多源评价之前, 能搞清楚评价前后所涉变量间的相关关系, 对有效地运用多源评价将大有裨益。

另外, 多源绩效评价不能直接引用, 需根据

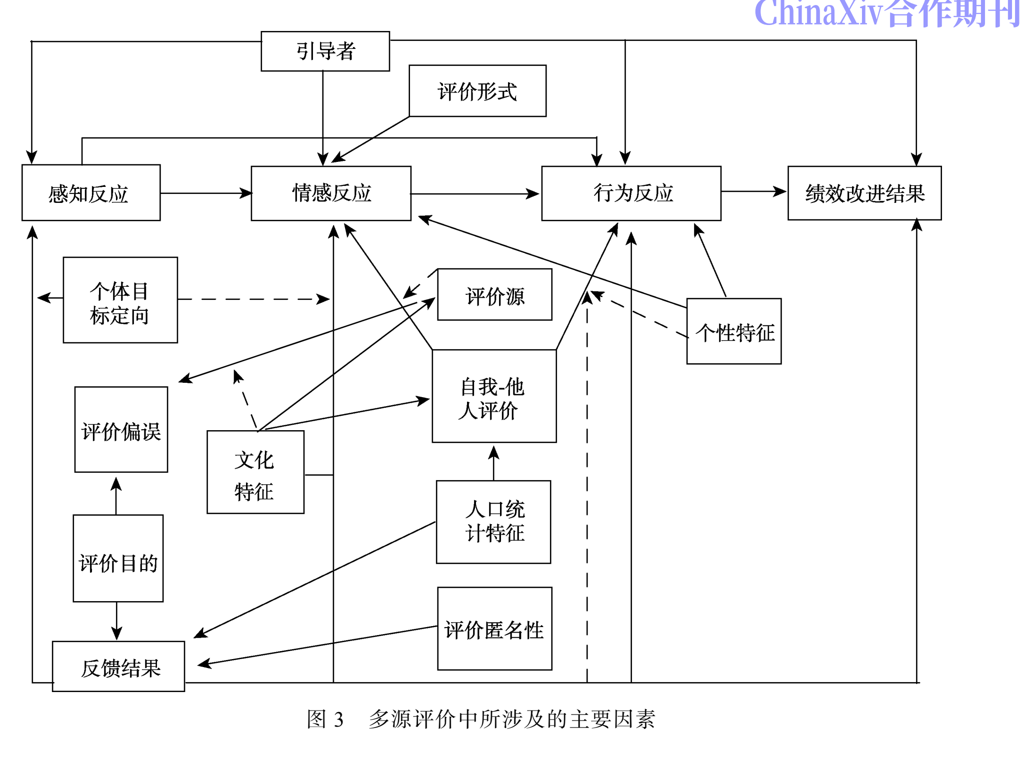


图3　多源评价中所涉及的主要因素

本国的文化特点等，进行相应的改进与完善(Harbi, Thursfield, & Bright, 2017)。因而，多源评价在我国企业中运用时，应注意与我国文化特点相结合。具体来说，我国的文化特征主要表现为，高权力距离、集体主义、不确定性规避程度较高，以及注重人际关系，即所谓的"关系文化"。在该文化背景下，较容易出现下级夸大对上级的绩效评价，同部门员工间相互评价被抬高，而压低对其他部门员工的评价等偏误。同时，较高的不确定性规避文化背景，也难以适应多源评价所需的高度开放性和互动性。为了避免上述评价偏误的发生，我国企业在运用多源评价时，应该制定严格且透明的绩效考核体系，避免"形式化"，如让员工明确考核目的及重要性，制定合理的考核指标，建立起组织的"反馈文化"，真正遵循公开、公正及公平原则。

6.2　未来研究展望

通过上述分析，发现多源评价所带来的绩效改进结果并不稳定。因而，未来的研究应从如何改进多源评价，以获得稳定的绩效改进结果方面，进一步开展相关的研究。

第一，关于如何提高评价多源评价过程的有效性与准确性问题，目前的研究还很少。绩效评价系统本身特点、员工个人特征及其与评价者间的关系，都影响着员工看待绩效评价的可视性、有效性及公平性(Sumelius, Björkman, Ehrnrooth, Mäkelä, & Smale, 2014)。当员工认为绩效评价过程不公平时，会表现出负面态度，如降低组织认同感(Farndale & Kelliher, 2013)，并影响其工作满意度(Yamazaki & Yoon, 2016)。因而，对多源评价过程的有效性进行评估与衡量，对于整个绩效评价的完善与改进有着重要影响。可参考对传统绩效评价有效性，所提出的相关标准及评价结构，如结合评价的目的性、公平性、准确性，以及该三大标准与被评价者反应间的关系，提出评估绩效评价效果的相关指标与方法(Iqbal, Akbar, & Budhwar, 2015; Ikramullah, van Prooijen, Lqbal, & Ul-Hassan, 2016) 。

第二，被评价者对评价结果的反应，对其后续的绩效改进行为影响较大。然而，关于个体对评价结果反应的研究中，学者们对被评价者的"反应"界定不清晰。基于此，本文将个体反应分为三类，并提出了"感知反应−情感反应−行为反应"的基本路径，但该分类及三者之间的关系是否准确，还需要未来作进一步验证与探讨。今后的研究可在本文的基础上，针对细分的"反应"类型，

对被评价者对多源评价结果的反应进行更深入的探讨, 通过改善被评价者对评价结果的反应情况, 从而促进其绩效改进行为。

最后, 如何对多源评价结果进行汇总, 也是重要的研究问题。多源评价过程中, 对不同的评价主体的评分, 大多都是以总平均的方式汇总, 或保持每类评价源的独立评分。然而, 不同的评价源的观察角度有差异, 应根据不同的评价指标, 予以不同的评价权重。比如, 在团队合作能力考核指标中, 同一团队的同级评价者与被评价者接触时间更多, 对被评价者在团队合作方面的能力相对更了解, 因而同级评价所占权重应最大。关于多源评价中不同评价源评分的赋权问题, 目前还少有文献讨论过。虽然 Markham 等(2014, 2015, 2017)提出不能对多源评价结果直接平均, 但其并没有对各评价源间如何赋权进行研究。对每类评价源给与科学合理的权重, 能给员工提供更准确的反馈结果, 从而促进其后续绩效改进行为。因而, 未来的研究应注重如何对评价结果进行科学有效的汇总, 以减少不同评价源带来的偏误。

参考文献

刘凤瑜. (2000). 一种有效的效绩评价模式: 360 度反馈的基本原理与过程. *南开管理评论*, (1), 71–74.

陆昌勤, 方俐洛, 凌文辁. (2001). 360 度反馈及其在人力资源管理中的效用. *中国管理科学*, *9*(3), 74–80.

许庆瑞, 王勇, 陈劲. (2002). 绩效评价源与多源评价. *科研管理*, *23*(3), 84–89.

Amundsen, S., & Martinsen, Ø. L. (2014). Self–other agreement in empowering leadership: Relationships with leader effectiveness and subordinates' job satisfaction and turnover intention. *The Leadership Quarterly*, *25*(4), 784–800.

Antonioni, D., & Park, H. (2001). The relationship between rater affect and three sources of 360-degree feedback ratings. *Journal of Management*, *27*(4), 479–495.

Atwater, L. E., & Brett, J. F. (2005). Antecedents and consequences of reactions to developmental 360° feedback. *Journal of Vocational Behavior*, *66*(3), 532–548.

Atwater, L. E., & Brett, J. F. (2006). 360-degree feedback to leaders: Does it relate to changes in employee attitudes? *Group & Organization Management*, *31*(5), 578–600.

Atwater, L. E, Brett, J. F., & Charles, A. C. (2007). Multisource feedback: Lessons learned and implications for practice. *Human Resource Management*, *46*(2), 285–307.

Atwater, L. E., Wang, M., Smither, J. W., & Fleenor, J. W. (2009). Are cultural characteristics associated with the relationship between self and others' ratings of leadership? *Journal of Applied Psychology*, *94*(4), 876–886.

Beenen, G., Pichler, S., & Levy, P. E. (2017). Self-determined feedback seeking: The role of perceived supervisor autonomy support. *Human Resource Management*, *56*(4), 555−569.

Bernardin, H. J., Tyler, C. L., & Villanova, P. (2009). Rating level and accuracy as a function of rater personality. *International Journal of Selection & Assessment*, *17*(3), 300–310.

Bernardin, H. J., Konopaske, R., & Hagan, C. M. (2012). A comparison of adverse impact levels based on top-down, multisource, and assessment center data: Promoting diversity and reducing legal challenges. *Human Resource Management*, *51*(3), 313–341.

Bernardin, H. J., Thomason, S., Buckley, M. R., & Kane, J. S. (2016). Rater rating-level bias and accuracy in performance appraisals: The impact of rater personality, performance management competence, and rater accountability. *Human Resource Management*, *55*(2), 321–340.

Beus, J. M., Jarrett, S. M., Bergman, M. E., & Payne, S. C. (2012). Perceptual equivalence of psychological climates within groups: When agreement indices do not agree. *Journal of Occupational and Organizational Psychology*, *85*(3), 454–471.

Bono, J. E, & Colbert, A. E. (2005). Understanding responses to multi-source feedback: The role of core self-evaluations. *Personnel Psychology*, *58*(1), 171–203.

Bono, J. E., Hooper, A. C., & Yoon, D. J. (2012). Impact of rater personality on transformational and transactional leadership ratings. *The Leadership Quarterly*, *23*(1), 132–145.

Braddy, P. W., Gooty, J., Fleenor, J. W., & Yammarino, F. J. (2014). Leader behaviors and career derailment potential: A multi-analytic method examination of rating source and self-other agreement. *The Leadership Quarterly*, *25*(2), 373–390.

Brett, J. F., & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology*, *86*(5), 930–942.

Brown, M., & Benson, J. (2003). Rated to exhaustion? Reactions to performance appraisal processes. *Industrial Relations Journal*, *34*(1), 67–81.

Brown, A., Inceoglu, I., & Lin, Y. (2017). Preventing rater biases in 360-degree feedback by forcing choice. *Organizational Research Methods*, *20*(1), 121–148.

Brutus, S., & Derayeh, M. (2002). Multisource assessment programs in organizations: An insider's perspective. *Human Resource Development Quarterly*, *13*(2), 187–202.

Carson, M. (2006). Saying it like it isn't: The pros and cons of 360-degree feedback. *Business Horizons, 49*(5), 395–402.

Chappelow, C. T. (2004). 360-degree feedback. In C. D. McCauley & E. van Velsor (Eds.), *The center for creative leadership handbook of leadership development* (pp. 58−84). San Francisco: Jossey-Bass.

Conway, J. M., Lombardo, K., & Sanders, K. C. (2001). A meta-analysis of incremental validity and nomological networks for subordinate and peer rating. *Human Performance, 14*(4), 267–303.

Culbertson, S. S., Henning, J. B., & Payne, S. C. (2013). Performance appraisal satisfaction. *Journal of Personnel Psychology, 12*(4), 189–195.

Eckert, R., Ekelund, B. Z., Gentry, W. A., & Dawson, J. F. (2010). I don't see me like you see me, but is that a problem? Cultural influences on rating discrepancy in 360-degree feedback instruments. *European Journal of Work and Organizational Psychology, 19*(3), 259–278.

Facteau, J. D., & Craig, S. B. (2001). Are performance appraisal ratings from different rating sources comparable? *Journal of Applied Psychology, 86*(2), 215–227.

Farndale, E., & Kelliher, C. (2013). Implementing performance appraisal: Exploring the employee experience. *Human Resource Management, 52*(6), 879–897.

Fleenor, J. W., Smither, J. W., Atwater, L. E., Braddy, P. W., & Sturm, R. E. (2010). Self–other rating agreement in leadership: A review. *The Leadership Quarterly, 21*(6), 1005–1034.

Harbi, S. A., Thursfield, D., & Bright, D. (2017). Culture, *Wasta* and perceptions of performance appraisal in Saudi Arabia. *The International Journal of Human Resource Management, 28*(19), 2792–2810.

Hassan, S., & Rohrbaugh, J. (2009). Incongruity in 360-degree feedback ratings and competing managerial values: Evidence from a public agency setting. *International Public Management Journal, 12*(4), 421–449.

Heslin, P. A., & Latham, G. P. (2004). The effect of upward feedback on managerial behavior. *Applied Psychology, 53*(1), 23–37.

Hoffman, B., Lance, C. E., Bynum, B., & Gentry, W. A. (2010). Rater source effects are alive and well after all. *Personnel Psychology, 63*(1), 119–151.

Ikramullah, M., Van Prooijen, J. W., Lqbal, M. Z., & Ul-Hassan, F. S. (2016). Effectiveness of performance appraisal: Developing a conceptual framework using competing values approach. *Personnel Review, 45*(2), 334–352.

Iqbal, M. Z., Akbar, S., & Budhwar, P. (2015). Effectiveness of performance appraisal: An integrated framework. *International Journal of Management Reviews, 17*(4), 510–533.

Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86*(1), 80–92.

Kampkötter, P. (2017). Performance appraisals and job satisfaction. *The International Journal of Human Resource Management, 28*(5), 750–774.

Karkoulian, S., Assaker, G., & Hallak, R. (2016). An empirical study of 360-degree feedback, organizational justice, and firm sustainability. *Journal of Business Research, 69*(5), 1862–1867.

Kim, K. Y., Atwater, L., Patel, P. C., & Smither, J. W. (2016). Multisource feedback, human capital, and the financial performance of organizations. *Journal of Applied Psychology, 101*(11), 1569–1584.

Kopperud, K. H., Martinsen, Ø., & Humborstad, S. I. W. (2014). Engaging leaders in the eyes of the beholder: On the relationship between transformational leadership, work engagement, service climate, and self-other agreement. *Journal of Leadership & Organizational Studies, 21*(1), 29–42.

Kossek, E. E., Huang, J. L., Piszczek, M. M., Fleenor, J. W., & Ruderman, M. (2017). Rating expatriate leader effectiveness in multisource feedback systems: Cultural distance and hierarchical effects. *Human Resource Management, 56*(1), 151–172.

Krasman, J. (2010). The feedback-seeking personality: Big five and feedback-seeking behavior. *Journal of Leadership & Organizational Studies, 17*(1), 18–32.

London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychology, 48*(4), 803–839.

London, M. (2001). The great debate: Should multisource feedback be used for administration or development only? In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback: The comprehensive resource for designing and implementing MSF processes* (pp. 368–385). San Francisco: Jossey-Bass.

London, M., & Smither, J. W. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review, 12*(1), 81–100.

Luffarelli, J., Gonçalves, D., & Stamatogiannakis, A. (2016). When feedback interventions backfire: Why higher performance feedback may result in lower self-perceived competence and satisfaction with performance. *Human Resource Management, 55*(4), 591–614.

Luthans, F., & Peterson, S. J. (2003). 360-degree feedback with systematic coaching: Empirical analysis suggests a winning combination. *Human Resource Management, 42*(3), 243–256.

Markham, S. E., Smith, J. W., Markham, I. S., & Braekkan, K. F. (2014). A new approach to analyzing the Achilles' heel of multisource feedback programs: Can we really trust ratings of leaders at the group level of analysis? *The Leadership Quarterly, 25*(6), 1120–1142.

Markham, S. E., Markham, I. S., & Smith, J. W. (2015). At the crux of dyadic leadership: Self-other agreement of leaders and direct reports—Analyzing 360-degree feedback. *The Leadership Quarterly, 26*(6), 958–977.

Markham, S. E., Markham, I. S., & Smith, J. W. (2017). A review, analysis, and extension of peer-leader feedback agreement: Contrasting group aggregate agreement vs. self-other agreement using entity analytics and visualization. *The Leadership Quarterly, 28*(1), 153–177.

Maylett, T. (2009). 360-degree feedback revisited: The transition from development to appraisal. *Compensation & Benefits Review, 41*(5), 52–59.

McCarthy, A. M., & Garavan, T. N. (2007). Understanding acceptance of multisource feedback for management development. *Personnel Review, 36*(6), 903–917.

Mero, N. P., Guidice, R. M., & Brownlee, A. L. (2007). Accountability in a performance appraisal context: The effect of audience and form of accounting on rater response and behavior. *Journal of Management, 33*(2), 223–252.

Mishra, V., & Roch, S. G. (2013). Cultural values and performance appraisal: Assessing the effects of rater self-construal on performance ratings. *The Journal of Psychology, 147*(4), 325–344.

Ng, K. Y., Koh, C., Ang, S., Kennedy, J. C., & Chan, K.-Y. (2011). Rating leniency and halo in multisource feedback ratings: Testing cultural assumptions of power distance and individualism-collectivism. *Journal of Applied Psychology, 96*(5), 1033–1044.

O'Neill, T. A., Goffin, R. D., & Gellatly, I. R. (2012). The use of random coefficient modeling for understanding and predicting job performance ratings: An application with field data. *Organizational Research Methods, 15*(3), 436–462.

Pichler, S. (2012). The social context of performance appraisal and appraisal reactions: A meta-analysis. *Human Resource Management, 51*(5), 709–732.

Pulakos, E. D., & O'Leary, R. S. (2011). Why is performance management broken? *Industrial and Organizational Psychology, 4*(2), 146–164.

Randall, R., & Sharples, D. (2012). The impact of rater agreeableness and rating context on the evaluation of poor performance. *Journal of Occupational and Organizational Psychology, 85*(1), 42–59.

Rogers, E., Rogers, C. W., & Metlay, W. (2002). Improving the payoff from 360-degree feedback. *Human Resource Planning, 25*(3), 44–54.

Rosen, C. C., Kacmar, K. M., Harris, K. J., Gavin, M. B., & Hochwarter, W. A. (2017). Workplace politics and performance appraisal: A two-study, multilevel field investigation. *Journal of Leadership & Organizational Studies, 24*(1), 20–38.

Russo, S. D., Miraglia, M., & Borgogni, L. (2017). Reducing organizational politics in performance appraisal: The role of coaching leaders for age-diverse employees. *Human Resource Management, 56*(5), 769−783.

Semeijn, J. H., Van Der Heijden, B. I. J. M., & Van Der Lee, A. (2014). Multisource ratings of managerial competencies and their predictive value for managerial and organizational effectiveness. *Human Resource Management, 53*(5), 773–794.

Shipper, F., Hoffman, R. C., & Rotondo, D. M. (2007). Does the 360 feedback process create actionable knowledge equally across cultures? *Academy of Management Learning & Education, 6*(1), 33–50.

Smither, J. W., & Walker, A. G. (2004). Are the characteristics of narrative comments related to improvement in multirater feedback ratings over time? *Journal of Applied Psychology, 89*(3), 575–581.

Smither, J. W, London, M., & Reilly, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology, 58*(1), 33–66.

Smither, J. W., Brett, J. F., & Atwater, L. E. (2008). What do leaders recall about their multisource feedback? *Journal of Leadership & Organizational Studies, 14*(3), 202–218.

Spence, J. R., & Keeping, L. M. (2010). The impact of non-performance information on ratings of job performance: A policy-capturing approach. *Journal of Organizational Behavior, 31*(4), 587–608.

Sumelius, J., Björkman, I., Ehrnrooth, M., Mäkelä, K., & Smale, A. (2014). What determines employee perceptions of HRM process features? The case of performance appraisal in MNC subsidiaries. *Human Resource Management, 53*(4), 569–592.

Vecchio, P. R., & Anderson, R. J. (2009). Agreement in self-other ratings of leader effectiveness: The role of demographics and personality. *International Journal of Selection and Assessment, 17*(2), 165–179.

Verbos, A. K., Miller, J. S, & Goswami, A. (2014). Employee social cognition and performance evaluation process reactions. *Personnel Review, 43*(4), 515–535.

Woehr, D. J., Sheehan, M. K., & Bennett, J. W. (2005). Assessing measurement equivalence across rating sources: A Multitrait-Multirater approach. *Journal of Applied Psychology, 90*(3), 592–600.

Wong, K. F. E., & Kwong, J. Y. Y. (2007). Effects of rater goals on rating patterns: Evidence from an experimental field study. *Journal of Applied Psychology, 92*(2), 577–585.

Wood, D., Harms, P., & Vazire, S. (2010). Perceiver effects as projective tests: What your perceptions of others say about you. *Journal of Personality and Social Psychology*, *99*(1), 174–190.

Xie, J. L., Chen, Z. G., & Roy, J. P. (2006). Cultural and personality determinants of leniency in self-rating among Chinese people. *Management & Organization Review, 2*(2), 181–207.

Yamazaki, Y., & Yoon, J. (2016). A cross-national study of fairness in Asia: How perceptions of a lack-of-group bias and transparency in the performance evaluation system relate to job satisfaction. *Human Resource Management, 55*(6), 1059–1077.

The analysis of characteristics and internal mechanisms of multisource feedback

ZHANG Yun[1]; WENG Qingxiong[2]

([1] *School of Business and Administration, Southwestern University of Finance and Economics, Chengdu* 611130, *China*)

([2] *School of Management, University of Science and Technology of China, Hefei* 230026, *China*)

Abstract: Although the multisource feedback in foreign enterprises has become mature, it is still underdeveloped in China. Based on extant studies, we discuss and analyze the characteristics and internal mechanism of multisource feedback through three constructs, including appraisal process, rater sources, and ratee. From the perspective of appraisal process, multisource feedback has various objectives, and it often preserves the anonymity of raters as well as managers should pay attention to reasonably using appraisal results. In terms of the rater sources, the level of rating agreement between different raters is low, and raters often make halo error and leniency effect easily. From the view of ratee, the reaction of individuals to the feedback results is impacted by personality, feedback signal, and self-others agreement, etc. Additionally, the studies also show the influence of multisource feedback on performance is unstable. Based on these findings, there is a need for further study on how to enhance the validity and accuracy of appraisal process, improve individuals' reaction, and summarize the results of multisource feedback.

Key words: multisource feedback; rating agreement; rating biases